

# PRV

PATENT- OCH REGISTRERINGSVERKET  
Patentavdelningen

J1046 U.S. PTO  
09/853883  
05/10/01  
  
#5

Intyg  
Certificate



Härmed intygas att bifogade kopior överensstämmer med de handlingar som ursprungligen ingivits till Patent- och registreringsverket i nedannämnda ansökan.

This is to certify that the annexed is a true copy of the documents as originally filed with the Patent- and Registration Office in connection with the following patent application.

(71) Sökande              Global IP Sound AB, Stockholm SE  
Applicant (s)

(21) Patentansökningsnummer    0001728-5  
Patent application number

(86) Ingivningsdatum          2000-05-10  
Date of filing

Stockholm, 2001-04-19

För Patent- och registreringsverket  
For the Patent- and Registration Office

Hjördis Segerlund

Avgift  
Fee      170:-

CERTIFIED COPY OF  
PRIORITY DOCUMENT

AWAPATENT AB

Kontor/Handläggare

Stockholm/Kenneth Ahrengart/KAT

GLOBAL IP SOUND AB

Ansökningsnr

Vår referens

SE-2008636

1

ENCODING AND DECODING OF A DIGITAL SIGNALTechnical field of the invention

The present invention relates to encoding of a digital signal and its blocks of digital samples for transmission over a packet switched network.

Consequently, the present invention further relates to decoding of a digital signal and its blocks of digital samples received from a packet switched network.

Technical background and prior art

Telephony over packet switched networks, such as IP (Internet Protocol) based networks (mainly the Internet or Intranet networks) has become increasingly attractive due to a number of features. These features include such things as relatively low operating costs, easy integration of new services, and one network for voice and data. The speech or audio signal in packet switched systems is converted into a digital signal, i.e. into a bitstream, which is divided in portions of suitable size in order to be transmitted in data packets over the packet switched network from a transmitter end to a receiver end.

Packet switched networks were originally designed for transmission of non-real-time data and voice transmissions over such networks causes some problems.

Data packets can be lost during transmission, as they can be deliberately discarded by the network due to congestion problems or transmission errors. In non-real-time applications this is not a problem since a lost packet can be retransmitted. However, retransmission is not a possible solution for real-time applications. A packet that arrives too late to a real-time application cannot be used to reconstruct the corresponding signal since this signal already has been, or should have been,

delivered to the receiving speaker. Therefore, a packet that arrives too late is equivalent to a lost packet.

One characteristic of an IP-network is that if a packet arrives the content of it is undamaged. An IP-packet has a header which includes a CRC (Cyclic Redundancy Check) field. The CRC is used to check if the content of the packet is undamaged. If the CRC indicates an error, the packet is discarded. In other words, bit errors do not exist, only packet losses.

The main problem with lost or delayed data packets is the introduction of distortion in the reconstructed speech or audio signal. The distortion results from the fact that signal segments conveyed by lost or delayed data packets cannot be reconstructed. The speech coders in use today were originally designed for circuit switched networks with error free channels or with channels having bit-error characteristics. Therefore, a problem with these speech coders is that they do not handle packet losses very well.

Considering what has been described above as well as other particulars of a packet switched network, there are problems connected with how to provide the same quality in telephony over packet switched networks as in ordinary telephony over circuit switched networks. In order to solve these problems, the characteristics of a packet switched network have to be taken into consideration.

In order to overcome the problems associated with lost or delayed data packets during real-time transmissions, it is suitable to introduce diversity for the transmission over the packet switched network.

Diversity is a method which increases robustness in transmission by spreading information in time (as in interleaving in mobile telephony) or over some physical entity (as when using multiple receiving antennas). In packet transmission, diversity is preferably introduced on a packet level by finding some way to create diversity between packets. The simplest way of creating diversity

in a packet switched network is to transmit the same packet payload twice in two different packets. In this way, a lost or delayed packet will not disturb the transmission of the payload information since another 5 packet with identical payload, most probably, will be received in due time. It is evident that transmission of information in a diversity system will require more bandwidth than transmission of information in a regular system.

10 Many of the diversity schemes or diversity systems in the prior art have the disadvantage that the transmission of a sound signal does not benefit from the additional bandwidth needed by the transmitted redundant information under normal operating conditions. Thus, for 15 most of the time, when there are no packet losses or delays, the additional bandwidth will merely be used for transmission of overhead information.

Since bandwidth most often is a limited resource, it would be desirable if a transmitted sound signal somehow 20 could benefit from the additional bandwidth required by a diversity system. Preferably, it would be desirable if the additional bandwidth could be used for improving the quality of the decoded sound signal at the receiving end.

In "Design of Multiple Description Scalar 25 Quantizers", V. A. Vaishampayan, IEEE Transactions on Information Theory, Vol. 39, No.3, May 1993, the use of multiple descriptions in a diversity system is disclosed. The encoder sends two different descriptions of the same source signal over two different channels, and the 30 decoder reconstructs the source signal based on information received from the channel(s) that are currently working. Thus, the quality of the reconstructed signal will be based on one description if only one channel is working. If both channels work, the reproduced 35 source signal will be based on two descriptions and higher quality will be obtained at the receiving end. In the article, the author addresses the problem of index

assignment in order to maximize the benefit of multiple descriptions in a diversity system.

In a system that transmits data over packet switched networks, one or more headers are added to each data 5 packet. These headers contain data fields with information about the destination of the packet, the sender address, the size of the data within the packet, as well as other packet transport related data fields. The size of the headers added to the packets constitutes 10 overhead information that must be taken into account. To keep the packet assembling delay of data packets small, the payload of the data packets have limited size. The payload is the information within a packet which is used by an application. The size of the payload, compared to 15 the size of the actually transmitted data packet with its included overhead information, is an important measure when considering the amount of available bandwidth. A problem with transmitting several relatively small data packets, is that the size of the headers will be substantial in comparison with the size of the 20 information which is useful for the application. In fact, the size of the headers will not seldom be greater than the size of the useful information.

To alleviate bandwidth problems, it is desirable to 25 reduce the bit rate by suitable coding of the information to be transmitted. One scheme frequently used is to code information data using predictions of the data. These predictions are generated based on previous information data of the same information signal. However, due to the 30 phenomenon that packets can be lost during transmission, it is not a good idea to insert dependencies between different packets. If a packet is lost and the reconstruction of a following information segment is dependent on the information contained in the lost 35 packet, then the reconstruction of the following information segment will suffer. It is important that this type of error propagation is avoided. Therefore the

ordinary way of using prediction to reduce the bit rate of a speech or audio signal is not efficient for this kind of transmission channels, since such prediction would lead to error propagation. Thus, there is a problem 5 in how to provide prediction in a packet switched system when transmitting data packets with voice or audio signal information.

The use of prediction is a common method in speech coding to improve coding efficiency, i.e. for decreasing 10 the bit rate. An example is the predictive coding technique for Differential PCM (DPCM) coders disclosed in "Digital Coding of Waveforms: Principles and Applications to Speech and Video", N.S. Jayant and P. Noll, Prentice Hall, ISBN 0-13-211913-7 01, 1984. The prediction of a 15 signal sample is computed by a predictor based on a previous quantized signal sample, i.e. the prediction is backward adaptive. The computed prediction sample is then subtracted from the original sample which is to be predicted. The result of the subtraction is the error 20 obtained when predicting the signal sample using the predictor. This resulting prediction error is then quantized and transmitted to a receiving end. At the receiver the prediction error is added to a regenerated 25 prediction signal from a predictor corresponding to the predictor at the transmitting end. This combination of the received prediction error with a calculated prediction value will enable a reconstruction of the original signal sample at the receiver end. This kind of coding leads to bit rate savings since redundancy is 30 removed and the prediction error signal has lower power than the original signal, so that less bits are needed for the quantization of the error signal at a given noise level.

As stated above, this kind of encoding/decoding of 35 speech or audio over a packet switched network leads to error propagation if a packet is lost. When a packet is not received, the prediction value calculated in the

decoder will be based on samples of the last packet that was received. This will result in a prediction value in the decoder that differs from the corresponding prediction value in the encoder. Thus, the received quantized prediction error will be added to the wrong prediction value in the decoder. Hence, a lost packet will lead to error propagation. If one would consider to reset the prediction state after each transmitted/received packet, there would be no error propagation. However, this would lead to a low quality of the decoded signal. The reason being that if the predictor state is set to zero, the result will be a low quality of the prediction value during encoding and, thus, the generation of a prediction error with more information content. This in turn will result in a low quality of the quantized signal with a high noise level since the quantizer is not adapted to quantize signals with such high information content.

If a diversity system is implemented based on multiple descriptions, the incorporation of prediction will face additional problems which are due to the fact that the sound signal has several representations. If the above described scheme for predictive encoding/decoding is used together with multiple description quantizers, one of two problems will be present. The problem will be dependent on how the predictors are utilized at the transmitting/receiving end.

If each of the multiple description quantizers at the receiving end were to feed independent prediction filters, the prediction value for each description would be independent of the arrival of the other multiple descriptions. However, with this solution the offset of the different encoded representations will be different between different independent predictor outputs. Thereby the regular spacing between representations from the multiple quantizers is lost, and with that the optimized

improvement from receiving multiple descriptions is also lost.

Alternatively, all multiple descriptions could be constructed from the same predictor, thereby maintaining the optimized improvement from receiving multiple descriptions. However, if this prediction is from a pre-defined representation, e.g. a best representation obtained from a merger of all descriptions, then synchronization of the decoder with the encoder is lost if one (or more) description of the multiple descriptions is not received due to a packet loss when transmitting that description from the encoder at the transmitting end to the decoder at the receiving end.

Thus, as stated above, there is a problem in how to use prediction for reducing the bit rate of a speech or audio signal for transmission over a packet network, since a lost packet with a signal information segment negatively will affect the reconstruction of the following signal information segment.

When using multiple descriptions, the transmission of the sound signal will require more bandwidth than if a single description was used. In such a system, it would be even more interesting to use prediction in order to reduce the required bandwidth. However, as described above, there is a problem in how to implement the predictive encoding/decoding mechanism in such a system, while maintaining the basic gain of multiple description quantization.

30 Summary of the invention

An object of the present invention is to overcome at least some of the above-mentioned problems of using predictive coding/decoding for reducing the bandwidth required when transmitting a digitized sound signal over a packet switched network.

According to the present invention, this object is achieved by the methods having the features as defined in the appended claims.

Thus, the present invention provides an advantageous way of encoding/decoding digital samples for transmission/reception over a packet switched network. This is performed by lossless encoding the digital samples, and lossless decoding of the corresponding code words, based on generated prediction samples. During encoding, the generated prediction samples are quantized separately from the quantization of the digital samples. The predictions are used in the index domain in the form of quantized indices during encoding/decoding of the digital signal.

One advantage of using predictions in this way is that the predictor can be configured to operate in the same way at the receiving end as at the transmitting end, and it will not be necessary to transmit any extra prediction information to the receiving end.

According to a first aspect of the invention, a method is provided for encoding a digital signal and its blocks of digital samples for transmission over a packet switched network, the method including the steps of: quantizing the digital samples; generating prediction samples based on previous quantized digital samples of said digital signal; and lossless encoding the quantized digital samples based on the generated prediction samples.

Consequently, according to a second aspect, a method is provided for decoding a digital signal and its blocks of digital samples received from a packet switched network, the method including the steps of: lossless decoding code words of received data packets into received quantization levels; generating prediction samples based on previously received quantized digital samples of said digital signal; deriving, based on the generated prediction samples, received quantized digital

samples of said digital signal from said quantization levels; and de-quantizing said received quantized digital samples into digital samples of said digital signal.

5 Predictions based on the quantized digital samples are generated either directly as quantization indices of prediction samples, or as samples which are quantized after its generation using the same set of quantization levels as used for the quantized digital samples, or a completely different set of quantization levels.

10 Thus, when basing the lossless encoding/decoding on prediction errors, these errors are generated in the quantization index domain as the difference between quantized prediction samples and quantized digital samples of the signal. This can be compared with  
15 traditional predictive encoding/decoding where the prediction errors are generated before quantization, in contrast to after the quantization. In the case of prediction before the quantizer, the quantizer quantizing the prediction errors will not be optimized for  
20 quantizing signals with a high information content.

Similarly, signals with high information content would be present if a predictor state is forced to zero in the beginning of a new block in order to avoid error propagation between different blocks of digital samples.  
25 In such a case the prediction error signal will basically be the original digital signal. However, with the solution according to the invention, the prediction error is used to enhance the performance of a lossless encoder.

Thus, using the present invention, a bad prediction value will still enable a good quality of the transmitted signal sample, the trade-off lies in that the bit savings of the predictive encoding/decoding will be low.

The predictor state of each predictor is advantageously set to zero when generating predictions samples during lossless encoding/decoding of a beginning of a block of digital samples. In general, the generation of a prediction sample during the lossless

encoding/decoding operation is based on one or more previous quantized digital samples. However, since a packet with a block of samples can be lost or delayed, it is desired not to use prediction over block boundaries in 5 order to avoid error propagation. Any edge effects due to the manipulation of the predictor state will be avoided since the samples of the digital signal itself is continuously quantized before being combined with quantized predictions. In contrast, basing the 10 quantization on the prediction errors only would lead to edge effects due to the great variation of the predictions when these are manipulated in the beginning of a block.

In an embodiment, the lossless encoding/decoding is 15 conditioned based on a quantized index of a generated prediction sample. The quantized index is used for selecting one out of several look-up tables with which quantized digital samples are losslessly encoded to code words, or code words are losslessly decoded to quantized 20 digital samples.

The quantized prediction, used to condition the lossless encoding/decoding, can be complemented by, e.g., a coarsely quantized estimate of the signal or prediction error variance, or other coarsely quantized features 25 extracted from the past of the signal. Thus, a number of features can be extracted from the past of the signal, be coarsely quantized, and then used to condition a lossless encoder or decoder. Hence, a lossless encoder/decoder can be independently optimized and used for each possible 30 combination of indexes from the quantization of the extracted features. Examples of useful features for the encoding of speech signals are: a quantized prediction; the quantizer index from not only one but from several previous samples in the signal; a quantized estimate of 35 signal or prediction-error variance; an estimate for the direction of the waveform; and/or a voiced/unvoiced classification.

Some of the above features can be extracted per sample or per block of samples in the encoder and transmitted as side-information. Waveform direction is an example of such a feature suitable for transmission as 5 side-information, e.g., by use of a high-dimensional block code. A voiced/unvoiced classification is another. The side-information results in a product code for the lossless encoding. The encoding of this product code can be made either sequentially or with analysis-by- 10 synthesis.

However, the advantage of the bit rate reduction by lossless encoding/decoding based on predictions is less significant, and the bandwidth still a problem, if a very large overhead in the form of a header is added to the 15 encoded information before transmission of the data packet. This problem will occur if multiple descriptions of the digital signal is used in order to obtain diversity, a problem which however is solved by the present invention.

20 Preferably, the encoder/decoder of the present invention is a multiple description encoder/decoder, i.e. an encoder/decoder which generates/receives at least two different descriptions of a digital signal. Thus, the multiple descriptions thereby provide multiple block 25 descriptions for each block of digital samples.

The invention provides diversity based on multiple descriptions by transmitting/receiving different individual block descriptions of the same block of digital samples in different data packets at different 30 time instances. This so called time diversity provided by the delay between the block descriptions is particularly advantageous when a time localized bottleneck occurs in the packet switched network, since the chance of receiving at least one of the block descriptions of a 35 certain block increases when the different block descriptions are transmitted at different points in time in different packets. Preferably a predefined time

interval between the transmissions of two individual block descriptions of the same block of digital samples is introduced.

Advantageously, block descriptions of different 5 descriptions of the digital signal and relating to different blocks of digital samples are grouped together in the same packet. At least two consecutive blocks are represented by individual block descriptions from different descriptions of the digital signal. This is 10 advantageous since it avoids the extra overhead required by the headers of the packets that transmit the different block descriptions for one and the same block of digital samples, while still only one block description of a specific block of digital samples is lost or delayed when 15 a packet is lost or delayed.

Advantageously, lossless encoding/decoding is performed for each different block description individually. This will reduce the bit rate needed for the multiple descriptions that are transmitted. 20 Furthermore, individual predictors of the same type are used for the different descriptions at the transmitting and the receiving end, respectively. This eliminates the problem of lost synchronization between an encoder and a decoder which otherwise can occur if a packet with a 25 block description is lost when using a single predictor for the lossless encoding/decoding at the transmitting/receiving end.

The invention is suitable for a digital signal consisting of a digitized sound signal, in which case a 30 block of digital samples corresponds to a sound segment of the digitized sound signal.

According to the invention the digital signal is optionally an n-bit PCM encoded digitized sound signal. Preferably a 64 kbit/s PCM signal in accordance with the 35 standard G.711. The n-bit PCM encoded signal description is transcoded by a multiple description encoder to at least two descriptions using fewer than n bits for its

representation, e.g. two  $(n-1)$ -bit representations, three  $(n-1)$ -bit representations or four  $(n-2)$ -bit representations. At the receiver end, a multiple description decoder transcodes the received descriptions back to a single n-bit PCM encoded sound signal. The transcoding corresponds to a translation between a code word of one description and respective code words of at least two different descriptions. By transcoding the PCM coded signal into multiple descriptions, there is no need to first decode and then recode the PCM coded signal to be able to provide multiple descriptions.

Thus, the invention enables the use of predictive coding/decoding when using multiple descriptions for transmitting a digital signal, such as a digitized sound signal, over a packet switched network.

It is to be understood that the term digital signal sample used herein is meant to be interpreted as either the actual sample or as any form of representation of the signal obtained or extracted from one or more of its samples. Also, a prediction sample is meant to be interpreted as either a prediction of an actual digital signal sample or as any form of prediction of a representation obtained or extracted from one or more of the digital signal samples. Finally, a quantization level of a digital sample is either the index or the value of a quantized digital sample.

#### Brief description of the drawings

Further features and advantages of the invention will become more readily apparent from the appended claims and the following detailed description of a number of exemplifying embodiments of the invention when taken in conjunction with the accompanying drawings in which like reference characters are used for like features, and wherein:

Fig. 1 shows one exemplifying way of realizing multiple descriptions in accordance with state of the art;

5 Fig. 2 shows an overview of the transmitting part of a system for transmission of sound over a packet switched network;

Fig. 3 shows an overview of the receiving part of a system for transmission of sound over a packet switched network;

10 Figs. 4a and 4b show overviews of a Sound Encoder at the transmitting part and of a Sound Decoder at the receiving part, respectively, of a system for transmission of sound over a packet switched network in accordance with an embodiment of the present invention;

15 Figs. 5a and 5b show overviews of a Sound Encoder at the transmitting part and of a Sound Decoder at the receiving part, respectively, of a system for transmission of sound over a packet switched network in accordance with another embodiment of the present

20 invention;

25 Figs. 6a and 6b show overviews of a Sound Encoder at the transmitting part and of a Sound Decoder at the receiving part, respectively, of a system for transmission of sound over a packet switched network in accordance with yet another embodiment of the present invention; and

30 Fig. 7 shows some of the element of the transmitting part of a system for transmission of sound over a packet switched network in accordance with a further embodiment of the present invention.

#### Detailed description of preferred embodiments

In Fig. 1, one exemplifying way of realizing multiple descriptions of a source signal, such as a sound signal, is illustrated. This approach is known in the art and is one example of multiple descriptions that can be used by the present invention. However, other suitable

ways of implementing multiple descriptions may equally well be used together with the present invention. In Fig. 1 the quantization levels of two different descriptions 100, 110 from two corresponding quantizers are shown. As 5 illustrated, both descriptions have the same quantization step size  $Q$ , but description 110 has quantization levels that are shifted with half of the quantization step size  $Q$  with respect to the quantization levels of description 100. From these two descriptions 100 and 110, a 10 combination leads to a combined description 120 with finer quantization step size  $Q/2$ . Using the two coarse quantizers, a bit rate of  $2R$  is required to match the performance of a single fine quantizer with bit rate  $R+1$ . For example, if each description 100 and 110 has 4 15 quantization levels, each will require 2 bits to code these levels, i.e. a total of 4 bits. If a finer quantizer would be used for the combined description 120, the 7 quantization levels would require 3 bits when coded. For high  $R$ , this will constitute a significant 20 increase of the bit rate when using two coarse quantizers for providing multiple descriptions instead of one finer quantizer providing a single description.

In Fig. 2 a block diagram of the transmitting part of a system for transmission of sound over a packet 25 switched network is shown. The sound is picked up by a microphone 210 to produce an analog electric signal 215, which is sampled and quantized into digital format by an A/D converter 220. The sampling rate of the sound signal is dependent on the source of the sound signal and the 30 desired quality. Typically, the sampling rate is 8 or 16kHz for speech signals, and up to 48kHz for audio signals. The quality of the digital signal is also affected by the accuracy of the quantizer of the A/D converter. For speech signals the accuracy is usually 35 between 8 and 16 bits per sample. In a typical system, the transmitting end includes a Sound Encoder 230 in order to compress the sampled digital signal further.

According to the present invention, an additional purpose of the Sound Encoder 230 is to modify the representation of the sound signal before transmission, with the intent to increase the robustness against packet losses and delays in the packet switched network. The sampled signal 225 is input to the Sound Encoder 230 which encodes the sampled signal and packetizes the obtained encoded signal into data packets. The data packets 235 are then transferred to a Controller 240 which adds sequencing and destination address information to the data packets, in order to make the packets suitable for transmission over a packet switched network. The data packets 245 are then transmitted over the packet switched network to a receiver end.

In Fig. 3 a block diagram of the receiving part of a system for transmission of sound over a packet switched network is shown. A Controller 350 receives data packets from the packet switched network, strips addressing information and places the data packets 355 in a Jitter buffer 360. The Jitter buffer 360 is a storage medium, typically RAM, which regulates the rate by which data packets 365 exit the Jitter buffer 360. The physical capacity of the jitter buffer is such that incoming data packets 355 can be stored. Data packets 365 which exit the Jitter buffer 360 are inputted to a Sound Decoder 370. The Sound Decoder 370 decodes the information in the data packets into reproduced samples of a digital sound signal. The digital signal 375 is then converted by a D/A-converter 380 into an analog electric signal 385, which analog signal drives a sound reproducing system 390, e.g. a loudspeaker, to produce sound at the receiver end.

The design and operation of the Sound Encoder and the Sound Decoder, in accordance with an embodiment of the invention, will now be described in greater detail with reference to Figs 4a and 4b. Apart from what is being described below with respect to the sound

encoding/decoding blocks, the overall operation correspond to that previously described with reference to Figs. 2 and 3.

In Fig. 4a, a Sound Encoder for encoding a digital signal at a transmitting end in accordance with an embodiment of the invention is shown. The Sound Encoder includes a first Quantizer 400, a De-quantizer 410, a Delay block 420, a Predictor 430, a second Quantizer 440 and a Lossless Encoder 450. The De-quantizer 410 and the second Quantizer 440 are depicted with dashed lines since they are not necessary elements of this embodiment. The use of these optional elements will be described later in an alternative embodiment.

Correspondingly, in Fig. 4b, a Sound Decoder for decoding a digital signal at a receiving end in accordance with an embodiment of the invention is shown. The Sound Decoder includes a Lossless Decoder 455, a Quantizer 470, a Predictor 480, a Delay block 490 and De-quantizers 460 and 463. The Quantizer 470 and the De-quantizer 463 are depicted with dashed lines since they are not necessary elements of this embodiment. The use of these optional elements will be described later in an alternative embodiment.

The purpose of performing lossless encoding/decoding by means of the Lossless Encoder 450 and the Lossless Decoder 455 is to find a less bit-consuming way to describe the data that is transmitted from the transmitting end to the receiving end without loosing any information. Lossless encoding uses statistical information about the input signal to reduce the average bit rate. This is, e.g., performed in such way that the code words are ordered in a table after how often they occur in the input signal. The most common code words are then represented with fewer bits than the rest of the code words. An example of a Lossless Encoder known in the art that uses this idea is the Huffman coder.

Lossless encoding only works well in networks without bit errors in the received data. The code words used in connection with lossless encoding are of different length, and if a bit error occurs it is not possible to know when a code word ends and a new begin. Thus, a single bit error does not only introduce an error in the decoding of the current code word, but in the whole block of data. When the packet switched network is an IP (Internet Protocol)-network, all damaged data packets are automatically discarded. Thus, in such a packet switched network there will be no bit errors in data packets received at the receiver end. Therefore, lossless encoding, such as scalar or block Huffman coding, are according to the invention suitable for use for independent compression of each of the coded blocks of digital samples which blocks together constitutes the digital signal.

The Lossless Encoder 450 and the Lossless Decoder 455 of the embodiment of Figs. 4a and 4b both includes a table which is created to include all possible code words and their bit representation. Table look-ups are performed to losslessly encode a block of digital samples quantized by the Quantizer 400 before being transmitted as code words over the packet network. Correspondingly, at the receiver end, the code words of an encoded block of quantized digital samples are losslessly decoded to quantized digital samples which then are de-quantized by De-quantizer 460 to a reconstructed original block of digital samples.

In Fig. 4a digital samples of a digital signal received from the A/D-converter are quantized by quantizer 400 into quantized digital samples. For each quantized digital sample a prediction sample is generated by Predictor 430 based on one or more previously quantized digital samples. The predictor 430 generates a quantization index for the prediction sample based on the quantization levels, i.e. quantization indices or

quantization values, for these previously quantized digital samples, which levels have been outputted by the Quantizer 400 and delayed by the Delay block 420. The quantization index of a prediction sample is then 5 transferred to a Subtractor 445 where it is subtracted from the quantization index of a current quantized digital sample outputted from the Quantizer 400. The result from the Subtractor 445, i.e. the difference between the quantization index of the prediction sample 10 and the quantization index of the current quantized digital sample, is transferred to the Lossless Encoder 450. The Lossless Encoder encodes the current quantized digital sample by using the index difference received from the Subtractor 445 as an entry in a look-up table 15 for outputting a corresponding code word. The code words of a complete encoded block of quantized digital samples are eventually assembled to a separate packet which is transferred to a Controller. Alternatively, each code word of an encoded block is collected by the Controller 20 and then assembled to a separate packet for the encoded block. The Controller adds header information before transmitting the data packet over a packet switched network.

In Fig. 4b the Sound Decoder corresponding to the 25 embodiment of Fig. 4a is shown. Packets with code words, or code words of disassembled packets, are received from a Jitter buffer by the Lossless Decoder 455. Each code word is used by the Lossless Decoder to select an entry 30 in a look-up table for outputting a corresponding index difference, which in turn corresponds to a quantized digital sample. For each quantized digital sample a prediction sample is generated by Predictor 480 based on one or more previous quantized digital samples. Predictor 480 at the receiving end is configured to operate in the 35 same way as Predictor 430 at the transmitting end. The configuration of these predictors is typically such that the predictor state is zero, or close to zero, when

generating prediction samples corresponding to the initial quantized digital samples of a digital signal. In the same way as at the transmitting end, predictor 480 generates a quantization index based on the quantization levels, i.e. quantization indices or quantization values, of previously quantized digital samples, which levels implicitly have been outputted by the Lossless Decoder 455 and delayed by the Delay block 490. The quantization index of the generated prediction sample is then 5 transferred to an Adder 465 where it is added to the index difference outputted from the Lossless Decoder 455. The result from the Adder 465, i.e. the sum of the quantization index of the prediction sample and the current index difference outputted, is transferred to the 10 De-quantizer 460 where it is de-quantized before being transferred to a D/A-converter.

15

In alternative embodiments, the Sound Encoder includes the De-quantizer 410 and/or the second Quantizer 440 as depicted in Fig. 4a. Correspondingly, the Sound 20 Decoder in accordance with these alternative embodiments includes the Quantizer 470 and/or the De-quantizer 463.

Using De-quantizers 410 and 463 quantization values 25 of quantized digital samples will be inputted to the Predictor 430 and 480 rather than quantization indices and the Predictors will generate prediction samples based on values rather than indices.

If the Predictors 430 and 480 do not include 30 quantization tables for outputting quantization levels, such as indices, of the generated prediction samples, the Sound Encoder/Decoder will include Quantizers 440, 470 for providing quantization levels, e.g. indices, of the generated prediction samples. In this way the Subtractor 445 and the Adder 465 will still be fed with the quantization levels of the prediction samples. Moreover, 35 using the Quantizers 440 and 470 it is ascertained that the quantization levels of the generated prediction samples will be valid levels belonging to a predefined

set of levels, and not levels falling between different valid quantization levels.

According to the invention, in order to avoid error propagation, a generated prediction sample corresponding to a digital sample of one block of digital samples should not be based on digital samples of a previous block. In accordance with an embodiment, this is achieved by setting a predictor state of Predictor 430 to zero before a new block with quantized digital samples is encoded. Correspondingly, in the Sound Decoder at the receiving end, the predictor state of Predictor 480 is set to zero before decoding a new block with quantized digital samples. As an alternative to setting the predictor state to zero, state information can be included in each block of digital samples, or, the encoding/decoding can follow a scheme which uses no or little state information when encoding/decoding the beginning of a block.

According to another embodiment of the invention, conditional lossless encoding/decoding is used at the transmitting/receiving end as indicated in Figs. 5a and 5b. The overall function is similar to that described above besides the use of the generated prediction samples and the configurations of the look-up tables. In the Sound Encoder at the transmitting end the prediction samples generated by a Predictor 530, optionally quantized by Quantizer 540 in accordance with what has previously been described, is used for selecting one out of several look-up tables with code words within a Conditional Lossless Encoder 550. The quantized level, such as the index, of the current quantized digital sample from Quantizer 500 is used to select a specific entry of the selected look-up table. The Conditional Lossless Encoder will then output a code word corresponding to this specific entry of the selected table and transfer it to the Controller. Alternatively, all code words representing a block of digital samples

are first assembled to a packet which then is transferred to the Controller.

Correspondingly, with reference to Fig. 5b, the generated prediction sample at the receiving end is used  
5 for selecting a look-up table, out of several tables, within a Conditional Lossless Decoder 555. A code word received from the Jitter buffer is used to address a specific entry of the selected table, after which a corresponding quantized digital sample is outputted for  
10 de-quantization by a De-quantizer 560.

Thus, the Sound Encoder/Decoder of the present invention is designed to reduce the bit rate needed when transmitting a digital signal over a packet switched network. The block of digital samples on which the Sound  
15 Encoder/Decoder operates are preferably sound segments with digitized sound samples.

The present invention is not optimized for any specific kind of predictor. However, for sound signals one choice of predictor is the one obtained by LPC analysis of the quantized signal, eventually refined with  
20 a long-term predictor as is well known for a person skilled in the art. Also non-linear predictors, such as the one defined by the oscillator model disclosed in "Time-Scale Modification of Speech Based on a Non-linear  
25 Oscillator Model", G. Kubin and W. B. Kleijn, in Proc. Int. Conf. Acoust. Speech Sign. Process., (Adelaide), pp. I453-I456, 1994, can be used in the encoding/decoding scheme of the present invention. If a basic oscillator modeling is used, the prediction will already be in the  
30 quantized domain, and the additional quantization of the prediction can be avoided.

According to the invention the Sound Encoder/Decoder is further designed to increase the robustness against packet losses and delays in the packet switched network.  
35 This design to increase the robustness relies on representing the sound signal, or any digital signal in the general case, with multiple descriptions. This design

is illustrated in Figs. 6a and 6b in accordance with an embodiment of the invention. Apart from what is being described below with respect to the sound encoding/decoding blocks, the overall operation 5 correspond to that previously described with reference to Figs. 2 and 3.

In Fig. 6a, the Sound Encoder 630 at the transmitting end includes a Multiple Description Encoder 610 and a Diversity Controller 620. Correspondingly, the 10 Sound Decoder 670 of Fig. 6b at the receiving end includes a Diversity Controller 650 and a Multiple Description Decoder 680.

Turning now to Fig. 6a, the Multiple Description Encoder 610 of the Sound Encoder 630 encodes a sampled sound signal 625 in two different ways, thereby obtaining 15 two different bitstream representations, i.e. two different descriptions, of the sound signal. As previously described, each description has its own set of quantization levels, achieved, e.g., by shifting the 20 quantization levels of one description with half a quantization step. Correspondingly, if three descriptions were to be provided, the quantization levels of the second description would be shifted with a third step with respect to the first description, and the third 25 description with a third step with respect to the second description. Thus, as indicated in Fig. 6a, the sound signal may be encoded using more than two descriptions without departing from the scope of the present invention. However, for ease of description, only two 30 signal descriptions will be used in the herein disclosed embodiments of the invention.

Each description provides a segment description of an encoded sound signal segment of the sound signal. The Multiple Description Encoder 610 generates each 35 description and its segment descriptions by lossless encoding of the digitized sound samples in accordance with what has previously been described with reference to

Fig. 4a. Thus, a respective set of all the elements shown in Fig. 4a will be present in a Multiple Description Encoder 410 referred to by Fig. 6a for each generated description. Correspondingly, a respective set of all the 5 elements shown in Fig. 4b will be present for each description used in a Multiple Description Decoder referred to by Fig. 4b.

In Fig. 6a, the different segment descriptions of the same sound segment are transferred in respective 10 packets to the Diversity Controller 620. In Fig. 6a, two descriptions have been indicated,  $D_1$  and  $D_2$ . The consecutive segments  $n$ ,  $n+1$ ,  $n+2$ , and so on, are represented by description  $D_1$  as segment descriptions  $D_1(n)$ ,  $D_1(n+1)$ ,  $D_1(n+2)$  . . . , which segment descriptions 15 are transferred in respective consecutive data packets 615, 616, 617 from the Multiple Description Encoder 610 to the Diversity Controller 620. Correspondingly, the same segments are also represented as segment descriptions  $D_2(n)$ ,  $D_2(n+1)$ ,  $D_2(n+2)$  . . . by description 20  $D_2$  and are also transferred in respective data packets to the Diversity Controller. Thus, each sound segment of the sound signal 625 is represented by one segment description of each description, e.g. in Fig. 6a sound segment  $n+1$  is represented by segment description  $D_1(n+1)$  25 of description  $D_1$  and by segment description  $D_2(n+1)$  of description  $D_2$ .

The Diversity Controller 620 dispatches the packets received from the Multiple Description Encoder 610 in accordance with the diversity scheme used. In Fig. 6a the 30 Diversity Controller 620 sequences each segment description of one sound segment in separate packets. The packets containing different segment descriptions of the same sound segment are transferred to the Controller 640 at different time instances. For example, as indicated in 35 Fig. 6a, the two segment descriptions  $D_1(n)$  and  $D_2(n)$  of sound segment  $n$  is delivered to the Controller 640 in separate packets 621 and 622 at different points of time

$t_1$  and  $t_2$ . Thus, a delay of  $t_2 - t_1$  is introduced to create time diversity. A typical delay  $t_2 - t_1$  that could be used, in connection with typical sound segment lengths of 20 ms, is 10 ms. Upon reception of a packet from the  
5 Diversity Controller 620, the Controller 640 formats the packet, such as adding sequencing and destination address information, for immediate transmission on the packet switched network. Thus, the Controller 640 adds a header, H, with information to each packet. In the case of IPv4  
10 transport using UDP (User Datagram Protocol) and RTP (Real Time Protocol), the header size is 320 bits. For a typical speech segment length of 20 ms, this leads to 320 bits per 20 ms, i.e. to 16 kbit/s for the headers of each description used. If each speech segment is represented  
15 by two segment descriptions, the headers of the packets transferring the segment descriptions will together require a bit rate of  $2 \times 16 = 32$  kbit/s. This can be compared to the bit rate of 64 kbit/s for standard PCM (Pulse Code Modulated) telephony. Consequently, the  
20 overhead bit rate will be 50% (32 divided with 64) of the payload rate.

As previously described with reference to Fig. 3, packets are received at the receiver end by a Controller 350. The Controller removes header information and  
25 transfers the packets to the Jitter buffer 360, which in turn transfers the packets to the Sound Decoder 370. Turning now to Fig. 6b, the Diversity Controller 650 of the Sound Decoder 670 receives the packets with the different segment descriptions from a jitter buffer. The  
30 Diversity Controller then schedules the different segment descriptions of the same sound segment for transfer to the Multiple Description Decoder 680 at the same time. Thus, as illustrated in the Fig. 6b, the Multiple Description Decoder 680 will e.g. receive both packets 671 and 672 with respective segment descriptions  $D_1(n)$  and  $D_2(n)$  of sound segment n at the same time, and then  
35 both packets 674 and 675 with respective segment

descriptions  $D_1(n+1)$  and  $D_2(n+1)$  of sound segment  $n+1$ , and so on. The Multiple Description Decoder 680 will for each sound segment extract the joint information from the different packets and decode the sound signal segment for transfer to a D/A-converter. If, e.g., segment description  $D_1(n)$  did not arrive at the receiver end, or arrived too late, the Diversity Controller 650 will only schedule  $D_2(n)$  (if two descriptions are used) to the Multiple Description Decoder 680, which then will decode sound segment  $n$  of the sound signal with adequate quality from the single segment description  $D_2(n)$  received.

In Fig. 7 another embodiment of the present invention is shown. This embodiment differs from the one previously described with reference to Figs. 6a and 6b with respect to the organization of segment descriptions in the packets transmitted by the packet switched network. Thus, the difference lies in the packet assembling/disassembling performed at the transmitting/receiving end by the Diversity Controller of the Sound Encoder/Decoder. This difference will now be described below.

As described with reference to Figs. 6a and 6b, the overhead resulting from the headers of the different packets transferring different segment descriptions of the same sound segment is quite extensive. To alleviate this, segment descriptions of different descriptions and relating to different sound segments are grouped together in the same packet before transmission of the packet over the packet switched network. As shown in Fig. 7 the Diversity Controller 720 of the Sound Encoder at the transmitting end groups two individual segment descriptions of two consecutive sound segments together in each packet. The two segment descriptions of a packet belong to respective descriptions of the sound signal. For example, one packet will contain segment description  $D_2(n-1)$  of sound segment  $n-1$  and segment description  $D_1(n)$  of sound segment  $n$ . The Controller 740 will as previously

described add header information to each packet before transmitting the packet including the two segment descriptions over the packet switched network.

Thus, just as in the embodiment of Fig. 6, the Diversity Controller 720 of this embodiment will sequence each segment description of a sound segment in separate packets, and, as in the embodiment of Fig. 6, the packets containing different segment descriptions of the same sound segment will be transferred to the Controller 740 at different time instances. In Fig. 7, the two segment descriptions  $D_2(n)$  and  $D_1(n+1)$  of sound segment  $n$  and  $n+1$  are delivered to the Controller 740 in packet 722. Thus, segment  $n+1$  must have been encoded before segment description  $D_2(n)$  can be transferred to the controller. Segment description  $D_1(n)$  on the other hand was transferred in a previous packet 721 to the controller. If a sound segment is 20 ms, the transfer of  $D_2(n)$  must be delayed with 20 ms compared with the transfer of  $D_1(n)$  since  $D_2(n)$  is to be scheduled in the same packet 722 as  $D_1(n+1)$ . Thus, this scheme will automatically provide time diversity since different segment descriptions of the same sound segment will be transferred to the Controller 740 with a 20 ms interval (given a sound segment length of 20 ms). Thus, in comparison with the embodiment of Fig. 6, an additional delay between the two segment descriptions of the same sound segment is automatically introduced with this scheme of assembling packets with several segment descriptions. This additional delay between segment descriptions provides an additional time diversity advantage and can be compensated for later in the transmission chain, e.g. by having lower delay settings in the jitter buffer at the receiving end.

Moreover, the amount of payload data in one packet according to this embodiment corresponds to the total amount of data generated from one sound segment,

therefore, the overhead information is not increased when creating time diversity with this scheme.

In correspondence with what has been described above, the Diversity Controller at the receiver end in 5 this embodiment will divide the received packets in their segment description parts before transferring the segment descriptions to the Multiple Description Decoder, in correspondence with what has been shown in Fig. 6b.

The effect of the time diversity scheme referred to 10 by Fig. 7 is again that if one packet is lost or delayed during transmission over the packet switched network, descriptions of all sound segments will still be available at the receiver end and no sound segment loss will be perceived.

According to an embodiment of the invention the Sound Encoder/Decoder encodes/decodes PCM indices of a standard 64 kbit/s PCM bitstream. This embodiment is for ease of description described by again referring to Figs 15 4a and 4b. As previously described the elements in respective Figs. 4a and 4b are present for each 20 description generated/decoded by the Sound Encoder/Decoder. However, the Quantizer 400 of Fig. 4a and De-quantizer 460 of Fig. 4b are exchanged with a respective Transcoder to be described below. Furthermore, 25 in case the digital signal is not already a PCM encoded signal, the Sound Encoder includes a PCM Encoder prior to its Transcoder and the Sound Decoder includes a PCM Decoder after its Transcoder. In this embodiment, the Sound Encoder again includes a Multiple Description 30 Encoder feeding a Diversity Controller with multiple descriptions of one and the same sound segment. Correspondingly, the Sound Decoder includes a Multiple Description Decoder receiving multiple descriptions of one and the same sound segment from a Diversity 35 Controller at the receiving end.

The Multiple Description Encoder of the Sound Encoder consists of an ordinary PCM Encoder followed by a

Transcoder. Thus, the digital signal received by the Sound Encoder from the A/D converter is encoded using an ordinary PCM Encoder. The obtained PCM bitstream is then transcoded, i.e. translated, into several bitstreams by 5 the Transcoder, after which each bitstream gives a coarse representation of the PCM signal. The corresponding Multiple Description Decoder at the receiving end includes a Transcoder for transcoding received multiple 10 bitstream descriptions to a single PCM bitstream. This PCM bitstream is then decoded by an ordinary PCM Decoder before being transferred to a D/A-converter. The method of transcoding, or translating is exemplified below where one 64 kbit/s PCM bitstream is transcoded into two 15 bitstreams which provide multiple descriptions of the PCM signal.

A standard 64 kbit/s PCM Encoder using  $\mu$ -law log compression encodes the samples using 8 bits/sample. This gives 256 different code words, but the quantizer only consists of 255 different levels. The zero-level is 20 represented by two different code words to simplify the implementation in hardware. According to the embodiment, each quantization level is represented by an integer index, starting with zero for the most negative level and up to 254 for the highest level. The first of the two 25 bitstreams is achieved by removing the least significant bit of each of the integer indices. This new index represents a quantization level in the first of the two coarse quantizers. The second bitstream is achieved by adding one to each index before removing the least 30 significant bit. Thus, two 7-bit representations are achieved from the original 8-bit PCM representation. Decoding of the two representations can either be performed on each individual representation, in case of packet loss, or on the two representations in which case 35 the original PCM signal is reconstructed. The decoding is simply a transcoding back into the PCM indices, followed by table look-up.

Alternatively, the PCM Encoder is a standard 64 kbit/s PCM Encoder using A-law log compression. In this case the number of levels in the quantizer is 256, which is one more than in a  $\mu$ -law coder. To represent these 256 levels using two new quantization grids, and be able to fully reconstruct the signal, one grid with 128 levels and one with 129 levels is needed. It would be desired to use two 7-bit grids like in the  $\mu$ -law case, however the problem with the extra quantization level has to be solved. According to the invention each quantization level is represented by an integer index, starting with zero for the most negative level and up to 255 for the highest level. The exact same rule as in the  $\mu$ -law case is used to form the new indices, except when representing index number 255. The index number 255 is represented with index number 126 for the first quantizer and index number 127 for the second instead of 128 and 127, which would be obtained if the rule would be followed. The decoder has to check this index representation when transcoding the two bitstreams into the A-law PCM bitstream. If only the first of the two descriptions is received after transmission, and the 255<sup>th</sup> index was encoded, the decoder will introduce a quantization error that is a little higher than for the other indices.

An encoded PCM signal includes a high degree of redundancy. Therefore, it is particularly advantageous to combine the use of PCM signals with lossless encoding/decoding of the multiple descriptions derived from a PCM signal.

If the digital signal received by the Sound Encoder already is represented as a 64 kbit/s PCM bitstream, and if the Sound Decoder at the receiving part should output a 64 kbit/s PCM bitstream, the PCM Encoder at the transmitting part and the PCM Decoder at the receiving part will not be needed. In this case the Multiple Description Encoder of the present invention receives the PCM bitstream and converts the PCM indices to the 0 to

254 representation described above. This representation  
is fed directly to the Transcoder, which transcodes the  
bitstream into two new bitstreams using the simple rules  
given above. At the receiver end of the system the  
5 information in the received packets are collected by the  
Diversity Controller. If all packets arrive the  
Transcoder merges and translates the information from the  
multiple descriptions back into the original PCM  
bitstream. If some packets are lost the original  
10 bitstream cannot be exactly reconstructed, but a good  
approximation is obtained from the descriptions that did  
arrive.

Although the invention has been described above by  
way of example with reference to different embodiments  
15 thereof, it will be appreciated that various  
modifications and changes can be made without departing  
from the scope of the invention as defined in the  
appended claims.

16  
17  
18  
19  
20  
21  
22  
23  
24

CLAIMS

1. A method of encoding a digital signal and its  
blocks of digital samples for transmission over a packet  
5 switched network, the method including the steps of:

quantizing the digital samples;  
generating prediction samples based on previous  
quantized digital samples of said digital signal; and  
lossless encoding the quantized digital samples  
10 based on the generated prediction samples.

2. The method as claimed in claim 1, wherein said  
step of generating prediction samples is preceded by de-  
quantization of the quantized digital samples, thereby  
15 obtaining the quantization values of said quantized  
digital samples.

3. The method as claimed in claim 1 or 2, including  
the step of quantizing the generated prediction samples,  
20 wherein said lossless encoding step is based on generated  
prediction samples having quantization levels of a  
predefined set of quantization levels.

4. The method as claimed in any one of claims 1 - 3,  
25 including the step of setting a state of a predictor  
generating said prediction samples to zero before  
starting to encode one of said blocks with digital  
samples.

30 5. The method as claimed in any one of claims 1 - 4,  
wherein said lossless encoding step is based on the  
quantization indices of said generated prediction  
samples.

35 6. The method as claimed in any one of claims 1 - 5,  
wherein said lossless encoding step for a specific,  
quantized digital sample includes outputting a specific

code word which corresponds to a specific entry of a table with code words, said specific entry being derived by means of a quantization index of a generated prediction sample corresponding to said specific 5 quantized digital sample.

7. The method as claimed in claim 6, wherein said specific entry is derived as the entry corresponding to the difference between the quantization index of said 10 specific quantized digital sample and said quantization index of said generated prediction sample.

8. The method as claimed in claim 6, wherein said table with code words is chosen among several tables with 15 code words based upon said quantization index of said generated prediction sample, wherein said specific entry is derived as the entry corresponding to said quantization index of said quantized digital sample.

20 9. The method as claimed in any one of claims 1 - 8, wherein said encoding is performed by a multiple description encoder, which multiple description encoder encodes each block of said blocks of digital samples by means of multiple block descriptions by performing the 25 steps of the encoding method individually for each generated block description.

10. The method as claimed in claim 9, including the additional step of transmitting, for each block of said 30 blocks of digital samples, at least two different block descriptions in respective data packets with a predefined time interval between the packets.

35 11. The method as claimed in claim 10, including grouping a respective block description of at least two different blocks of digital samples together for transmission in one and the same data packet.

12. The method as claimed in any one of claims 9 -  
11, wherein said digital signal is a digitized sound  
signal and said blocks of digital samples are sound  
5 segments, and wherein the encoding method in said  
multiple description encoder includes an initial step of  
transcoding an n-bit PCM represented digitized sound  
signal to at least two representations represented by  
fewer than n bits each and with respective sets of  
10 quantization levels for the segment descriptions of the  
sound segments of said digitized sound signal.

13. The method as claimed in any one of claims 1 -  
11, wherein said digital signal is a digitized sound  
15 signal and said blocks of digital samples are sound  
segments.

14. A method of decoding a digital signal and its  
blocks of digital samples received from a packet switched  
20 network, the method including the steps of:  
    lossless decoding code words of received data  
    packets into received quantization levels;  
    generating prediction samples based on previously  
    received quantized digital samples of said digital  
25 signal;  
    deriving, based on the generated prediction samples,  
    received quantized digital samples of said digital signal  
    from said quantization levels; and  
    de-quantizing said received quantized digital  
30 samples into digital samples of said digital signal.

15. The method as claimed in claim 14, wherein said  
step of generating prediction samples is preceded by de-  
quantization of the received quantized digital samples,  
35 thereby obtaining the quantization values of said  
quantized digital samples.

16. The method as claimed in claim 14 or 15,  
including the step of quantizing the generated prediction  
samples, wherein said deriving step is based on generated  
prediction samples having quantization levels of a  
5 predefined set of quantization levels.

17. The method as claimed in any one of claims 14 -  
16, including the step of setting a state of a predictor  
generating said prediction samples to zero before  
10 starting to decode one of said blocks with digital  
samples.

18. The method as claimed in any one of claims 14 -  
17, wherein said deriving step is based on the  
15 quantization indices of said generated prediction  
samples.

19. The method as claimed in any one of claims 14 -  
18, wherein said lossless decoding step for a specific  
20 quantized digital sample includes outputting a specific  
quantization level which corresponds to a specific entry  
of a table with quantization levels, said specific entry  
being selected by means of a received code word  
corresponding to said specific quantized digital sample.  
25

20. The method as claimed in claim 19, wherein said  
deriving step includes adding the quantization index of a  
generated prediction sample corresponding to said  
specific quantized digital sample to the quantization  
30 index of said specific quantization level.

21. The method as claimed in claim 19, wherein said  
table with quantization levels is chosen among several  
tables with quantization levels based upon a generated  
35 prediction sample corresponding to said specific  
quantized digital sample.

22. The method as claimed in any one of claims 14 -  
21, wherein said decoding is performed by a multiple  
description decoder, which multiple description decoder  
decodes each block of said blocks of digital samples  
5 based on at least two different received block  
descriptions by performing the steps of the decoding  
method preceding the de-quantizing step individually for  
each received block description.

10 23. The method as claimed in any one of claims 14 -  
22, including the steps of:

waiting a predefined time period for reception of at  
least two different packets including different block  
descriptions of one and the same block of digital  
15 samples;

performing the steps of the decoding method  
preceding the de-quantizing step with respect to those,  
one or several, different block descriptions of said  
block of digital samples received within said predefined  
20 time period; and  
de-quantizing the one, or a merger of the several,  
block descriptions.

24. The method as claimed in claim 23, wherein each  
25 received packet includes several block descriptions of  
several different blocks of digital samples grouped  
together, the method including the step of dividing  
successively received packets with respect to the  
included block descriptions, thereby obtaining several  
30 different block descriptions for each block of digital  
samples to be decoded.

25. The method as claimed in any one of claims 23 -  
24, wherein said digital signal is a digitized sound  
35 signal and said blocks of digital samples are sound  
segments, and wherein said digitized sound signal is a  
PCM encoded bitstream, and wherein any merger of said de-

quantizing step involves transcoding at least two segment representations, each represented by fewer than n bits, to a single n-bit PCM representation of said one and the same sound signal segment.

5

26. The method as claimed in any one of claims 14 - 24, wherein said digital signal is a digitized sound signal and said blocks of digital samples are sound segments.

10

27. A computer readable medium having computer executable instructions for causing a digital signal and its blocks of digital samples to be encoded for transmission over a packet switched network, the computer executable instructions performing the steps of the method as claimed in any one of claims 1 - 13.

28. A computer readable medium having computer executable instructions for causing a digital signal and its blocks of digital samples received from a packet switched network to be decoded, the computer executable instructions performing the steps of the method as claimed in any one of claims 14 - 26.

25

1  
2  
3  
4  
5  
6  
7  
8  
9

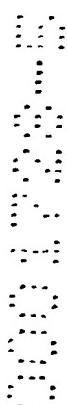
ABSTRACT

The invention relates to methods for encoding/decoding of a digital signal which is transmitted over a packet switched network. Prediction samples are generated at the transmitting and receiving end. The digital signal is lossless encoded at the transmitting end, and lossless decoded at the receiving end, based on the quantizations of generated prediction samples. During encoding, the generated prediction samples are quantized separately from the quantization of the digital samples. The predictions are used in the index domain in the form of quantized indices during encoding/decoding of the digital signal.

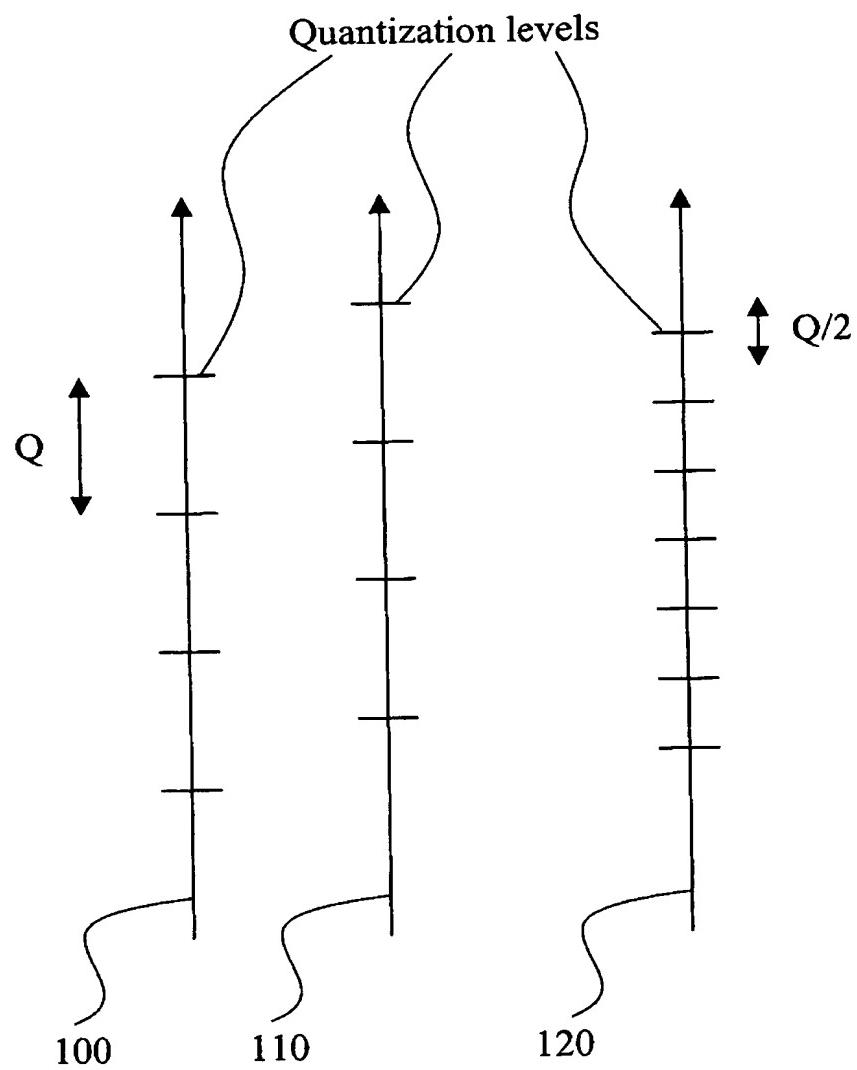
15

Fig. 4a.

20



1/8

**FIG. 1**

2/8

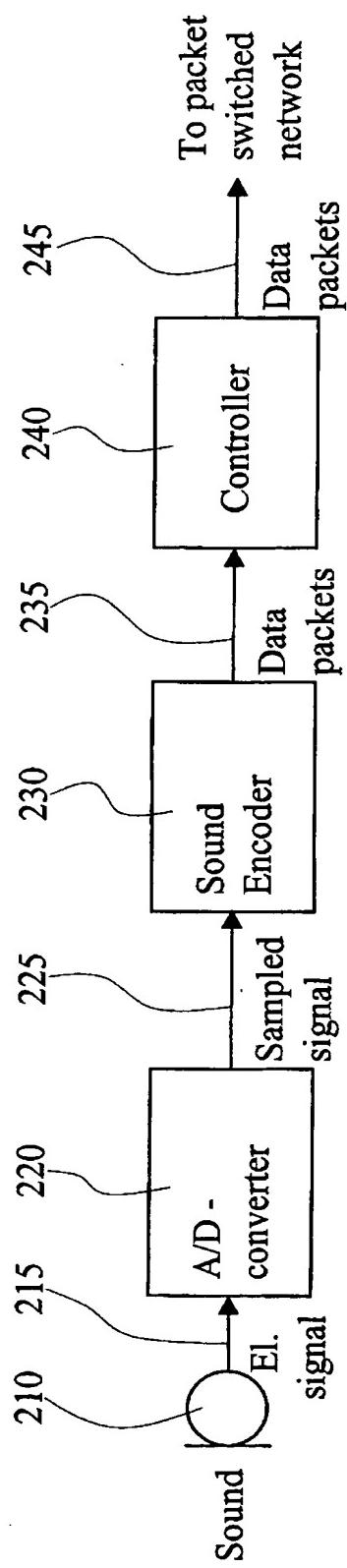


FIG. 2

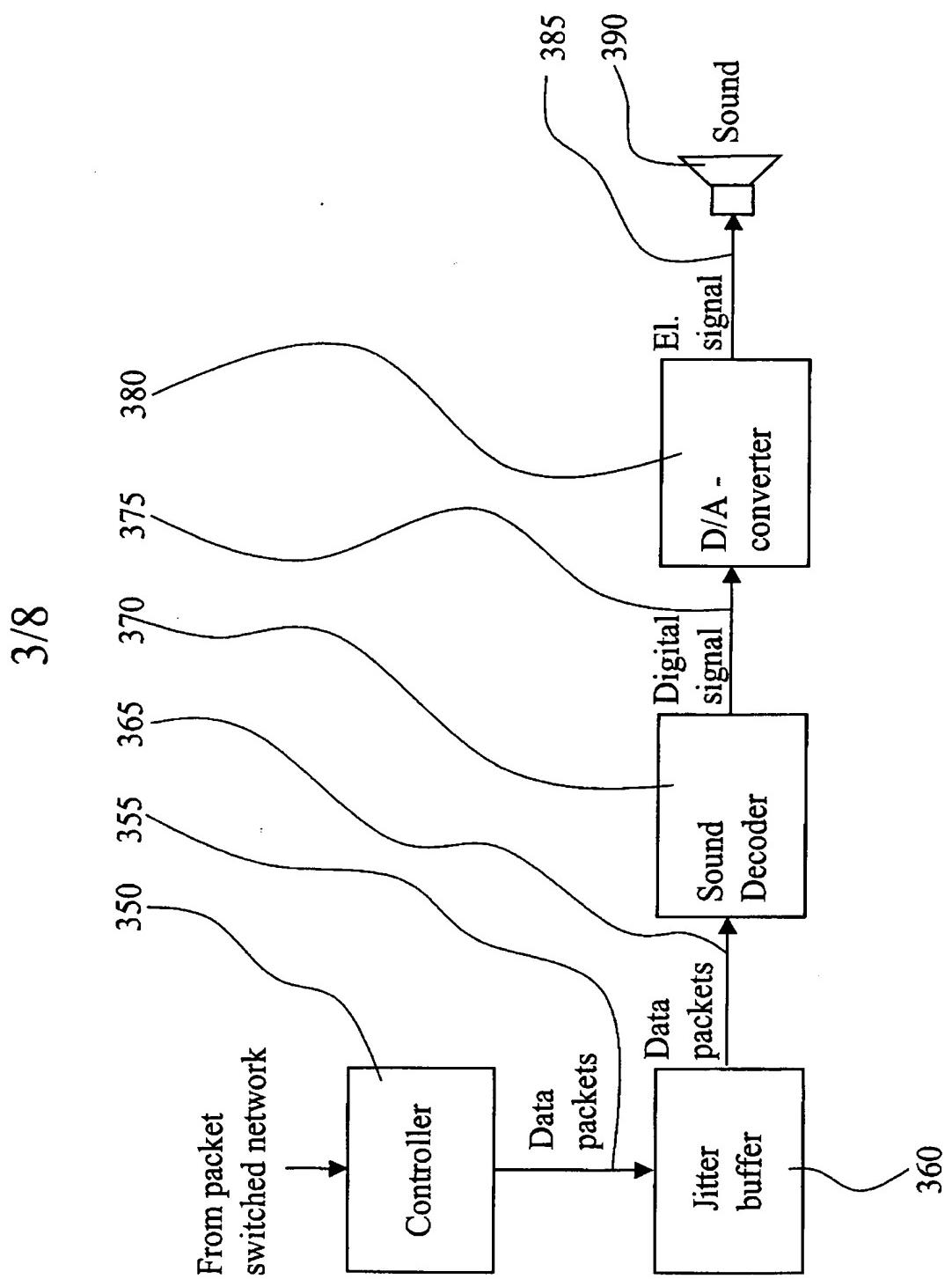


FIG. 3

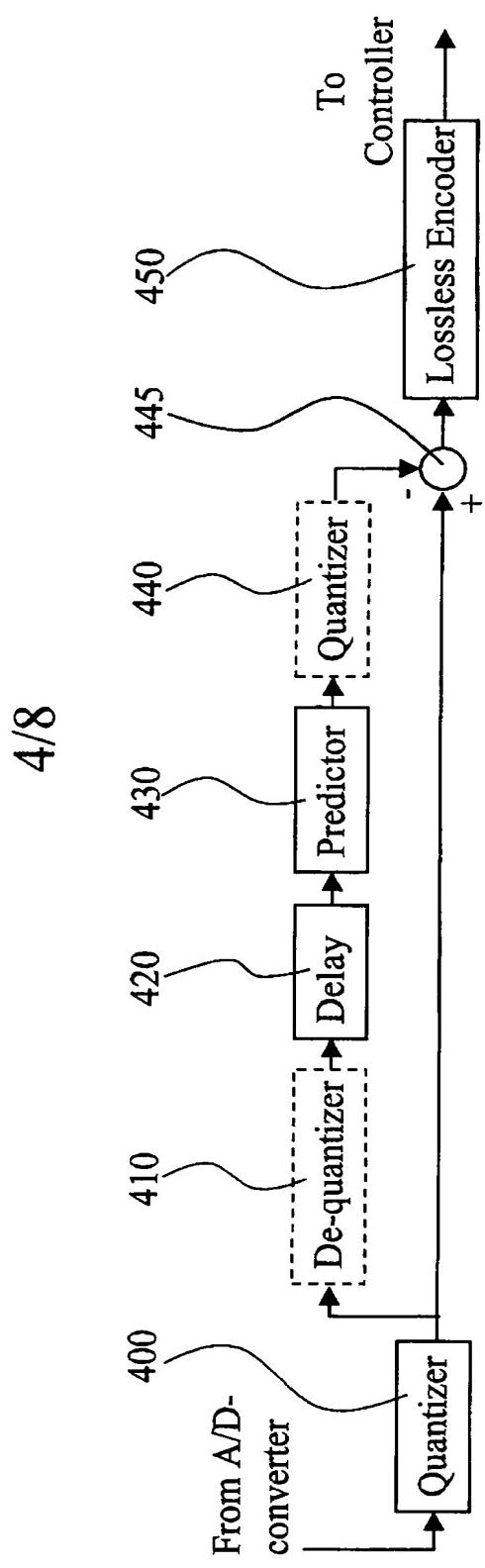


FIG. 4a

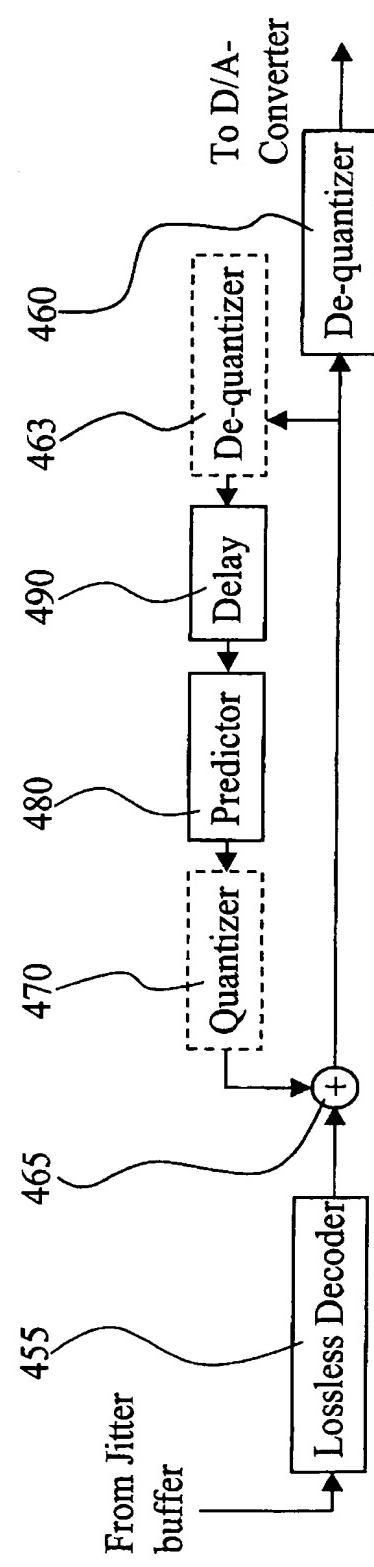


FIG. 4b

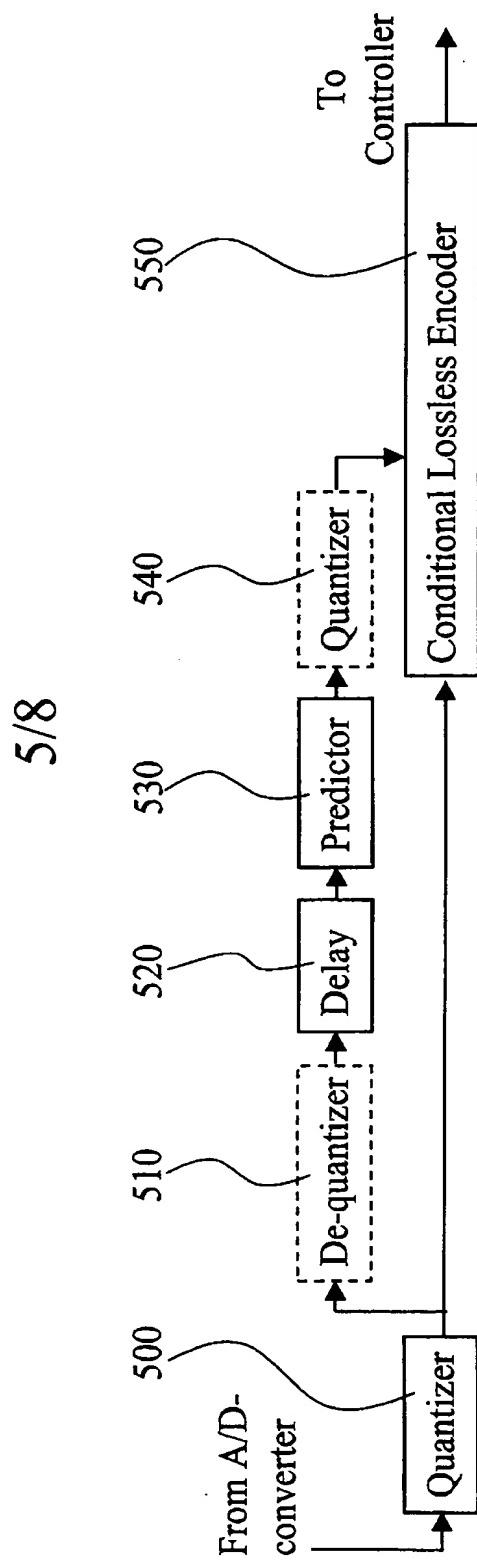


FIG. 5a

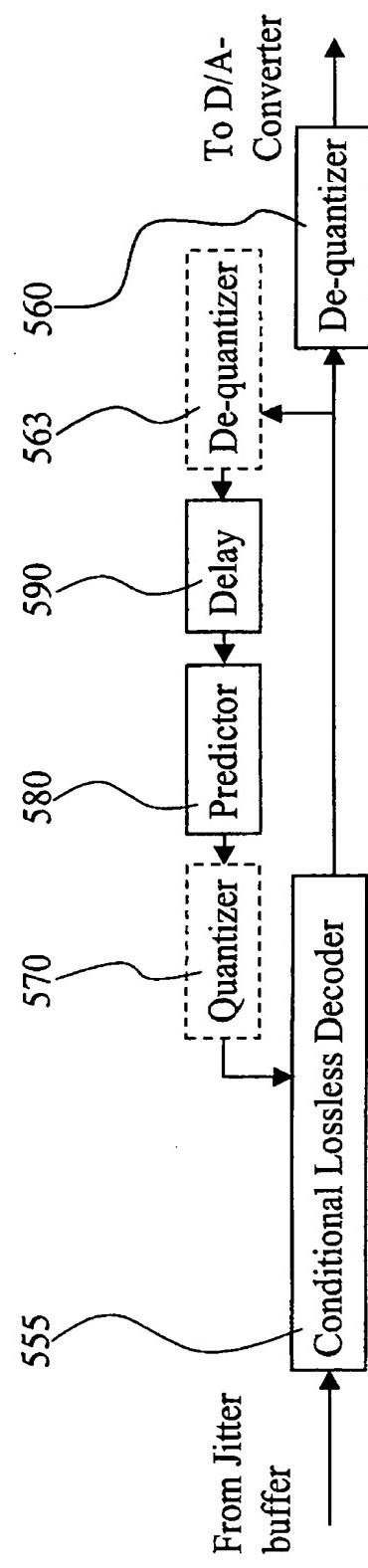
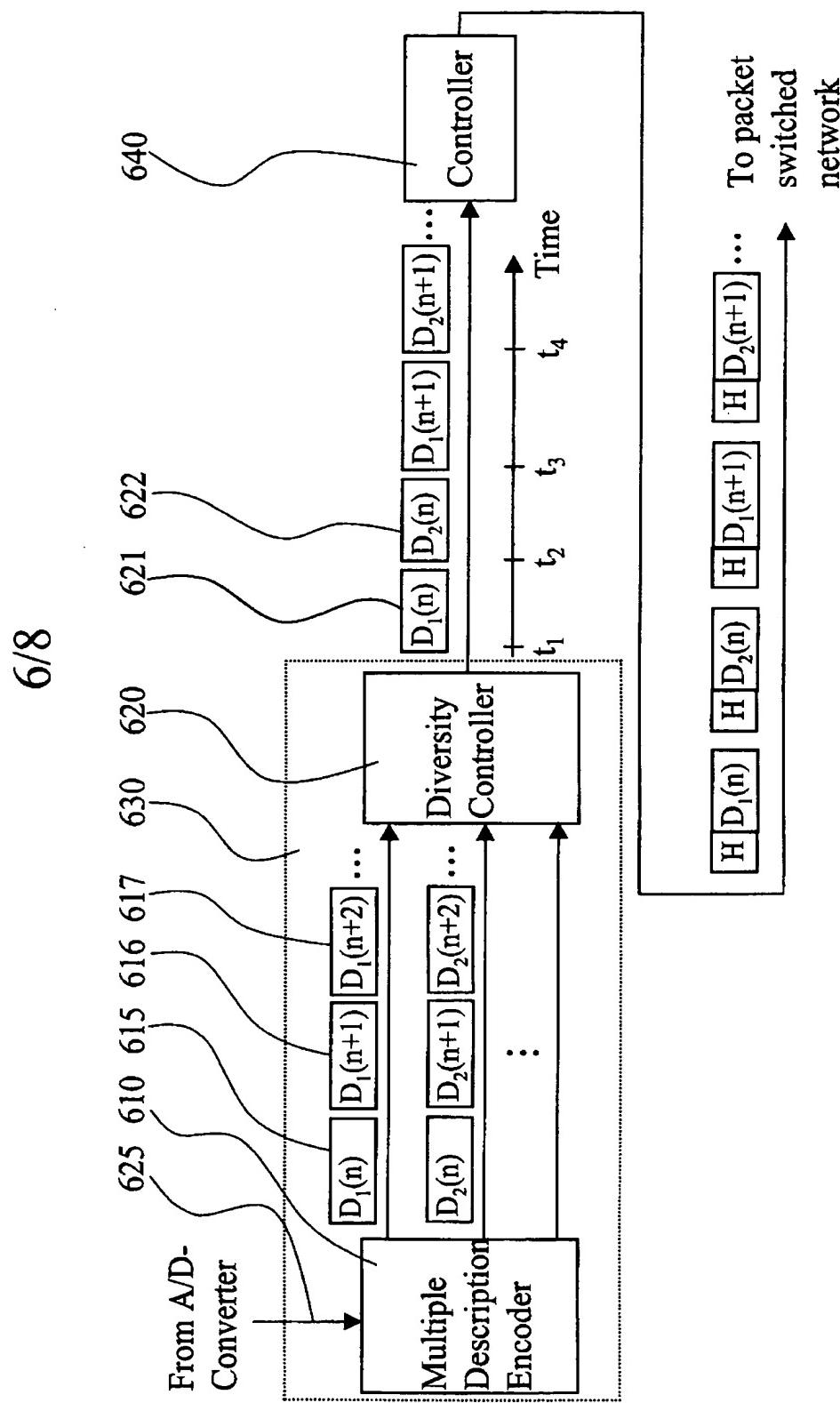
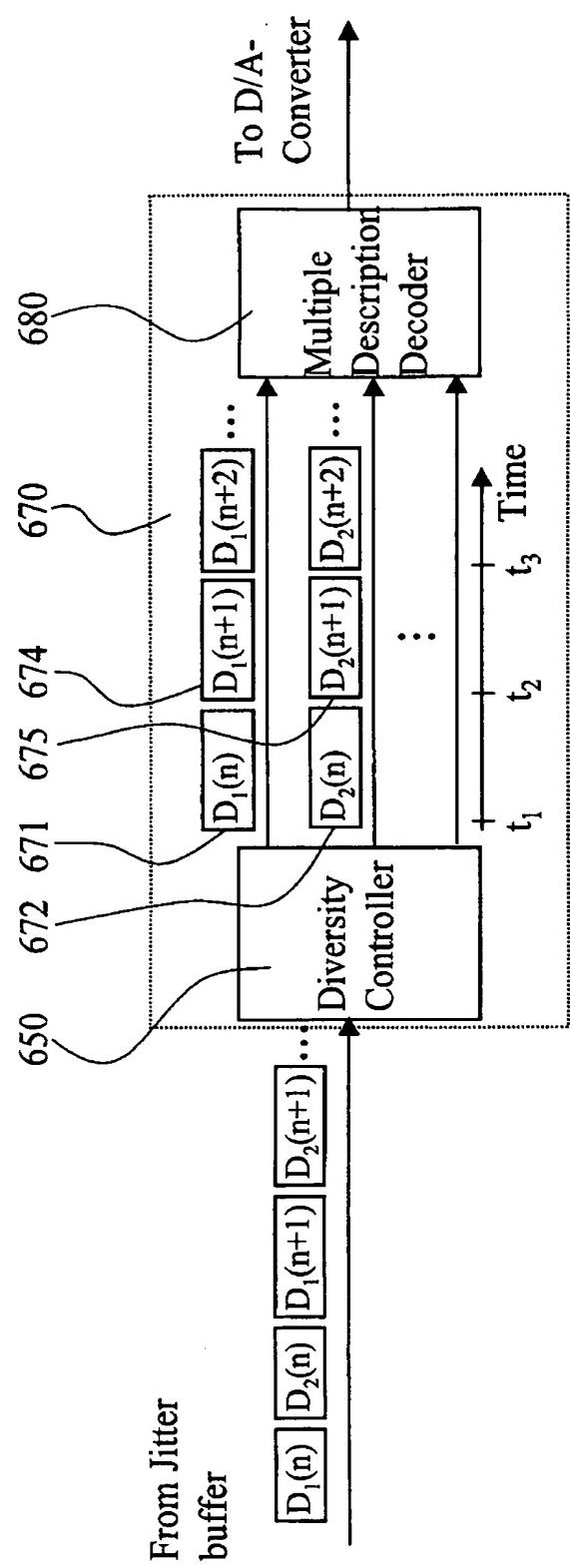


FIG. 5b

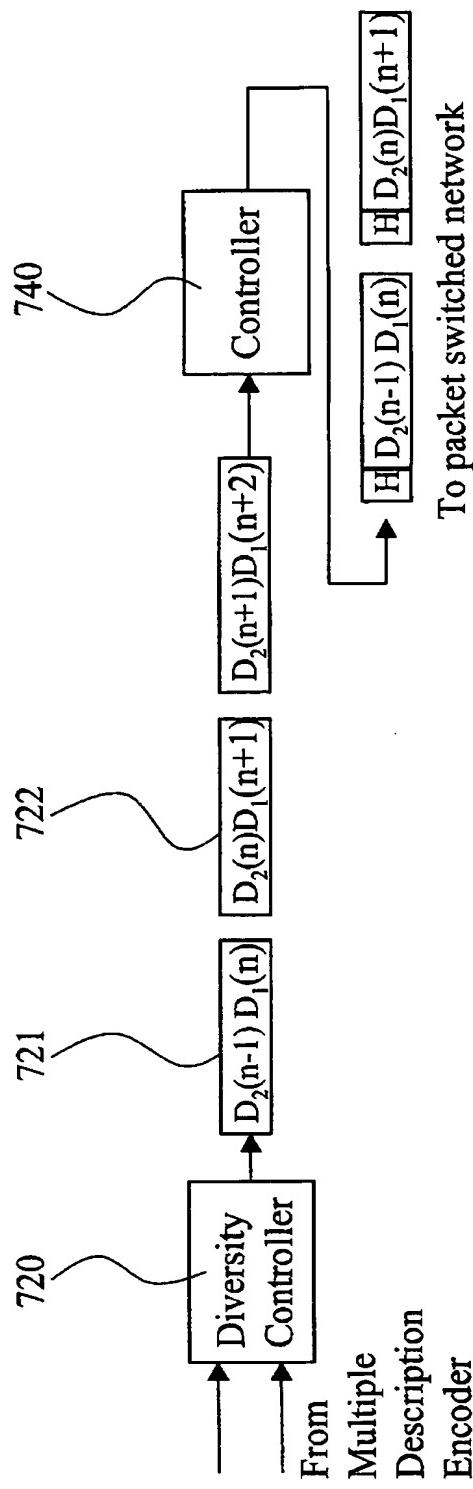
**FIG. 6a**

7/8

**FIG. 6b**

20047235

8/8

**FIG. 7**